Cognition 138 (2015) 35–48

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Verbal framing of statistical evidence drives children's preference inferences

Laura E. Garvin *, Amanda L. Woodward

Department of Psychology, University of Chicago, 5848 S. University Ave, Chicago, IL 60637, USA

ARTICLE INFO

Article history:
Received 11 November 2013
Revised 21 January 2015
Accepted 25 January 2015

Keywords:
Statistical reasoning
Theory of mind
Preferences
Bayesian reasoning
Conceptual development
Social cognitive development

ABSTRACT

Although research has shown that statistical information can support children's inferences about specific psychological causes of others' behavior, previous work leaves open the question of how children interpret statistical information in more ambiguous situations. The current studies investigated the effect of specific verbal framing information on children's ability to infer mental states from statistical regularities in behavior. We found that preschool children inferred others' preferences from their statistically non-random choices only when they were provided with verbal information placing the person's behavior in a specifically preference-related context, not when the behavior was presented in a non-mentalistic action context or an intentional choice context. Furthermore, verbal framing information showed some evidence of supporting children's mental state inferences even from more ambiguous statistical data. These results highlight the role that specific, relevant framing information can play in supporting children's ability to derive novel insights from statistical information.

1. Introduction

Reasoning about other people's minds, commonly called "theory of mind," can be challenging for young children and develops throughout early childhood (see Wellman & Liu, 2004). An important question in social cognitive development is how children begin to use information from others' behavior to make inferences about their mental states. In recent years, an influx of evidence has suggested that young children may be able to track the statistical patterns and regularities in a person's actions to help them learn about mental states – specifically, by using a person's choices to predict his or her preferences (Diesendruck, Salzer, Kushnir, & Xu, 2013; Kushnir, Xu, & Wellman, 2010; Lucas et al., 2014; Ma & Xu, 2011). One proposed mechanism for this ability is Bayesian reasoning, which formalizes the interaction between prior knowledge and statistical input in shaping the development of children's theories about the world. In layman's terms, Bayesian models propose that children interpret the patterns and regularities they experience in relation to the likelihood of those patterns occurring under various hypotheses (see Perfors, Tenenbaum, Griffiths, & Xu, 2011). Children begin with one or multiple theories about why certain events might occur ("priors" in Bayesian parlance), evaluate any incoming evidence in relation to how likely it is under each theory, and then adjust their theories appropriately.

Recently, several theorists have turned to Bayesian reasoning as a possible mechanism for how children may develop mature theory of mind capabilities (e.g., Gopnik, 2012; Gopnik & Wellman, 2012; Lucas et al., 2014; Xu & Kushnir, 2013). The promise of Bayesian reasoning is that through combining prior knowledge with statistical regularity information observed in the world, children may be able to derive *novel* insights, beyond what they

* Corresponding author at: Department of Psychology, 5848 S. University Ave, Chicago, IL 60637, USA. Tel.: +1 773 834 9791.
E-mail address: garvin@uchicago.edu (L.E. Garvin).

http://dx.doi.org/10.1016/j.cognition.2015.01.011
0010-0277/

had previously understood. Bayesian reasoning opens the door for a mechanism by which children may be able to rationally construct large-scale conceptual changes from data. As Gopnik and Wellman (2012) propose, "One attraction is that, at least in principle, this kind of inference would allow children to move from one structured hypothesis to another very different hypothesis based on patterns of evidence. Children need not merely fiddle with the details of an innately determined structure or simply accumulate more and more evidence. They could genuinely learn something new" (p. 1088).

Deriving novel insights by rationally evaluating patterns in the world requires children to have two distinct sets of skills. First, children must have sufficient statistical reasoning ability to be able to analyze statistical patterns in the world and interpret that analysis in relation to their prior hypotheses. In addition, children must also be able to choose appropriate hypotheses to consider and then intelligently use the statistical conclusions they have drawn to develop appropriate new hypotheses, a process that over many iterations may lead children to derive novel inferences. Most of the research exploring children's use of Bayesian reasoning to make mental state inferences has focused on the first set of abilities, and the evidence shows that young children indeed have the capacity to reason about a given mental state hypothesis in a Bayesian way (Kushnir et al., 2010; Lucas et al., 2014; Ma & Xu, 2011). In these studies, children were given prior information that clearly marked the relationship between the inference and the data. In other words, when children know what question they are trying to answer, they are able to use statistical information intelligently to answer this question. In this work, we focus on the second set of abilities: how children begin to determine which hypotheses to entertain and learn to generate new hypotheses and insights from the evidence they observe.

1.1. Statistical inference ability

Much of the previous work on Bayesian theories of social cognitive development has focused on whether children can interpret statistical patterns of evidence in an intelligent, rational way when given appropriate and relevant prior information. Evidence shows that they can. Beginning in infancy, children demonstrate a sophisticated understanding of the statistics of random sampling. Infants as young as 6 months of age are able to use the contents of a box to predict the probable identities of objects randomly selected from it (Denison, Reed, & Xu, 2013), and infants as young as 8 months can use a randomly selected sample to predict likely features of the population from which the sample was selected (Xu & Garcia, 2008; see also Denison, Konopczynski, Garcia, & Xu, 2006; Denison & Xu, 2009; Gweon, Tenenbaum, & Schulz, 2010; Xu & Denison, 2009). Infants can also use statistical patterns to make some inferences about intentional agents. Nine- to 12-month-old infants infer that ordered, regular patterns of objects are likely to have been caused by agents, and by 11 months of age, infants recognize that agents can intentionally violate statistical randomness (Ma & Xu, 2013; Newman, Keil, Kuhlmeier, & Wynn, 2010; Xu & Denison, 2009). Fifteen-month-old infants can also use an agent's intentional, statistically non-random sampling to make predictions both about the sampling process and about how the sample's features may generalize to the population (Gweon et al., 2010). As these studies demonstrate, infants possess abilities that may be the building blocks of more mature statistical reasoning skills.

When given appropriate prior information, preschool children and older toddlers can use their statistical reasoning abilities to support broader inferences about the mental causes of actions. Recent studies have demonstrated this by investigating children's inferences about an agent's preferences by analyzing the statistical regularities in the agent's choices (Kushnir et al., 2010; Lucas et al., 2014; Ma & Xu, 2011). These studies have shown that statistical information can influence children's preference inferences. Kushnir et al. (2010) found that when a puppet repeatedly chose an uncommon toy from a box containing many of one kind of toy and very few of another, young children inferred that the puppet purposefully chose that uncommon toy because he liked it (see also Diesendruck et al., 2013). Similarly, Ma and Xu (2011) found that when an experimenter selected several boring toys from a box containing mostly interesting toys, 2-year-old children predicted that the experimenter liked the boring toys. These studies provide reliable evidence that young children have the capacity to reason about behavioral information in a rational way in order to make inferences about specific mental states (see Lucas et al., 2014).

It is important to note that in all previous studies investigating children's ability to infer preferences from statistically regular patterns in choice behavior, children were provided with contextual information highlighting the relationship between the statistical evidence they were about to observe and the mental state inference they would be asked to make. In Bayesian parlance, children were guided to a specific prior hypothesis that they could test using the statistical information provided. In these studies, the contextual information came in the form of explicit verbal framing information establishing that the agent's behavior reflects her mental states (Diesendruck et al., 2013; Kushnir et al., 2010; Ma & Xu, 2011). That is, the agent's choices were described using the same mental states terms that were used in later test questions, giving children a relevant context to draw upon when observing and interpreting the agent's choices. For example, in the studies by Kushnir et al. (2010) and Diesendruck et al. (2013), children were told when introduced to a puppet that it "really likes" some toys then later were asked to predict which toys the puppet likes. Therefore, as Bayesian models would predict, children in previous studies may have taken advantage of the contextual framing of the agent's actions to help drive their interpretations of the statistical data and in turn, their inferences about the agent's preferences.

1.2. Selecting appropriate hypotheses

The work done to date on children's preference inferences is important because it shows that even when provided with adequate prior information about a possible

mental state driving certain behavior, children only infer that the agent's preference is the specific mental cause of his or her behavior when the statistical information supports this inference (Kushnir et al., 2010; Ma & Xu, 2011). However, previous work leaves open the question of how children determine which mental states may be possible causes of behavior, and importantly, whether children can make novel mental state inferences from observing behavior in more ambiguous contexts. How do children arrive at a "very different hypothesis" about mental states from what they had previously considered, and can they truly "genuinely learn something new" from analyzing patterns of behavior? (Gopnik & Wellman, 2012).

Recent evidence has shown that in the domain of causal reasoning, in which children must reason about which entities were likely to have caused certain physical outcomes, verbal framing information may be one important method for guiding children's interpretations of statistical information. When 4-year-old children were provided with verbal background information about a causal problem, they were able to analyze statistical evidence to infer the correct causal mechanism regardless of the difficulty of the inferential task; however, when children were not provided with this additional verbal contextual support, they succeeded only when the inferential task was simple, failing at more challenging tasks (Butler & Markman, 2012a). Similarly, young children's causal inferences are more accurate when the relevant statistical events are described using causal compared to non-causal language (Muentener & Schulz, 2012), and in ambiguous statistical situations, children make different causal inferences about identical statistical data when they are provided with verbal and observational evidence that certain items are more or less likely to be the causal mechanism (Griffiths, Sobel, Tenenbaum, & Gopnik, 2011; Sobel, Tenenbaum, & Gopnik, 2004). As these studies demonstrate, when children are faced with difficult inferential tasks, their conclusions about statistical information can be greatly affected by having a specific verbal framework to help guide them to an appropriate possible casual mechanism, thus shaping their interpretations.

1.3. The current studies

In the current studies, we investigated children's inferences about mental states when they were not necessarily provided with clear information about which mental state may have caused an agent's behavior. We did this by varying the verbal contextual information provided before and during the study to make the potential mental state driving the behavior more or less salient. We adapted methods from Kushnir et al. (2010; see also Diesendruck et al., 2013) to assess preschool children's ability to infer others' preferences from their statistically non-random choices when given verbal framework information that is either specifically related to preferences and mental states or that contains little to no mental state information. If preschool children are able to posit novel psychological causes of behavior even in ambiguous situations, then they may be able to successfully infer the character's preferences regardless of the verbal framework provided. If, however, preschool children require contextual information that is precisely relevant to the task at hand, then they may only make accurate preference inferences when the verbal framework is specifically related to preferences and not when it is less clearly related to mental states. Through this work, we can build on previous research showing that children have the ability to rationally analyze statistical information in light of a specific hypothesis and explore whether preschool children can also select and evaluate appropriate hypotheses and develop novel inferences in more ambiguous situations.

To investigate these questions, we showed preschool-age children an event in which a puppet intentionally selected five identical toys out of a box containing 17% of one kind of toy and 83% of another kind. In the first study, we explored the effects of specific verbal framework information on children's mental state inferences from statistically non-random behavior by showing two groups of children identical demonstrations of the puppet intentionally choosing the uncommon (17%) toy from the box, varying only whether this behavior was framed in terms of preference-related emotions or in terms of non-mentalistic actions. In a second study, we investigated whether specific verbal framework information affects children's judgments even in statistically ambiguous situations by showing children the puppet choosing toys in an intentional but plausibly random way – selecting the more common (83%) toy from the box. Finally, in a third study we investigated how children's inferences are affected by verbal framing information that highlights the intentionality of the puppet's choice but is not related to a specific mental state.

2. Study 1

Although previous work has demonstrated that young children can infer an agent's preferences based on her statistically non-random patterns of behavior, the information presented in these studies was always accompanied by contextual information highlighting the relationship between the agent's choices and her preferences (Kushnir et al., 2010; Ma & Xu, 2011). In this study, we investigated the importance of having verbal framing information that is specifically relevant to the task at hand by providing two groups of children with identical statistically non-random choice behavior and varying only whether these choices were verbally framed in terms of preferences or simply in terms of actions. Because the statistical information was equally informative in both conditions, any differences in responses must be due to the specific verbal context in which this information was presented.

2.1. Methods

2.1.1. Participants

Participants were 32 preschool-age children (range = 3 years, 6 months – 4 years, 0 months; mean = 3 years, 9 months; 16 females). Six additional children were tested but excluded because of child interference during the puppet's choices (2), failure to complete the study (2),

uncodable responses (1), or experimenter error (1). Participants were recruited from a large Midwestern city and tested either in an on-campus laboratory or at a local science museum. The sample of children was 59% Caucasian, 13% Hispanic, 13% multiracial, 6% Asian, 6% unreported, and 3% African-American.

2.1.2. Demonstration phase

Participants sat at a table across from an experimenter. The experimenter showed the child a small tray containing one example of each of the four kinds of small toys used in the experiment: a rubber frog, a rubber duck, a baseball, and a basketball, all approximately five to six centimeters in size, and asked the child to identify each of the toys, ensuring that the child's attention was drawn to each of the toys. The experimenter then removed the toys and introduced the child to a gorilla hand puppet that the experimenter operated, named Gabby or Gary to match the gender of the child.

During the puppet's introduction, the experimenter gave children different verbal framing information based on their randomly assigned condition. To control for the pedagogical intent and salience of the experimenter's utterances, the verbal information provided in each condition was designed to be approximately equal in duration and syntax. In the emotion framing condition, the experimenter introduced the puppet by saying to the child, "This is Gabby and she's going to play with some toys today. Some toys make her really *happy*, and other toys don't make her *happy*. Do you think you can help me figure out which toys make her *happy*?" Thus, children in this condition were told that the puppet has certain emotions about some toys, and they were specifically given information related to the questions they would be asked to answer later. In the action framing condition, the experimenter said, "This is Gabby and she's going to play with some toys today. Some toys she always *gets*, and other toys she doesn't *get*. Do you think you can help me watch and see what she *gets*?" The puppet's actions in this condition were described using purely behavioral terms, in contrast to the mental state terms used in the emotion framing condition.

The experimenter then put the puppet away and brought out a clear plastic box ($14 \times 14 \times 33$ cm) containing one of two sets of small toys, either basketballs and baseballs or ducks and frogs. The box of ducks and frogs contained eight of one kind of animal and 40 of the other, and the box of basketballs and baseballs, which were slightly larger toys, contained seven of one kind of ball and 35 of the other. For both toy sets, the toys were in a 1:5 ratio, meaning the box contained approximately 17% of one kind of toy and 83% of the other. Participants were counterbalanced on which set of toys they saw in the first block of the study (ducks and frogs or basketballs and baseballs) and which toy in the set was the more common kind of toy in the box. After introducing the puppet, the experimenter showed the child the contents of the box and asked the child to identify the toys inside.

For children in both conditions, the experimenter then brought out the puppet and said to it, "Hi Gabby! Are you going to play with some toys?" and the puppet nodded its assent. In the emotion framing condition the experimenter then said to the puppet, "I have some toys here and I think some might make you *happy* and I think some might not make you *happy*," and in the action framing condition, the experimenter said, "I have some toys here and I think some you might *get* and I think some you might not *get*." The puppet then chose five of the less common (17%) toys from the box, one at a time, and set them beside the box (see Fig. 1). After the puppet's choice, the experimenter said in the emotion framing condition, "Gabby, is that what you *want*?" and in the action framing condition, "Gabby, is that what you *got*?" Again, in this way, children in the emotion framing condition received information and scaffolding from the experimenter about the puppet's emotions and mental states, while children in the action framing condition did not receive this information. The

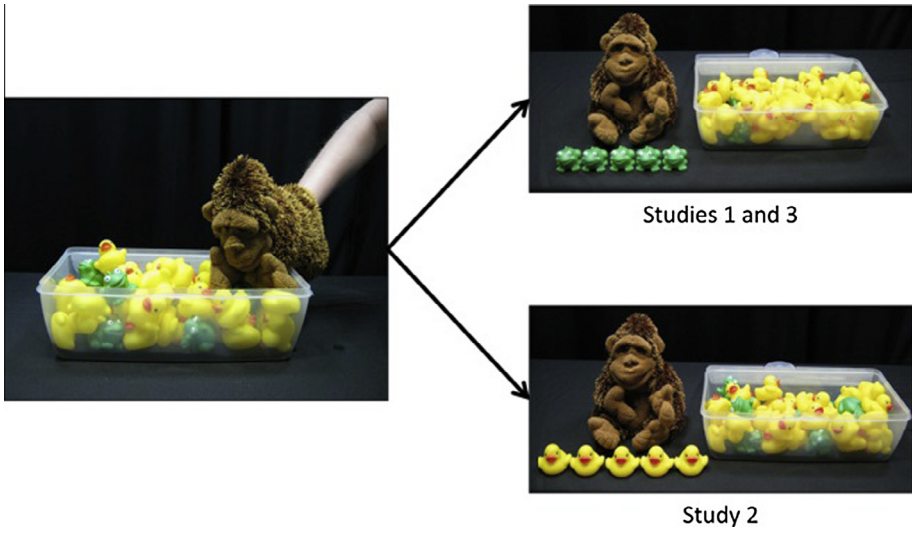

Fig. 1. Experimental set-up. In all studies, a gorilla puppet selected five of either the less common (Studies 1 and 3, top right) or more common (Study 2, bottom right) toys from a box.

puppet then played with the toys briefly by putting them all in a row. During the entirety of the demonstration, including the puppet's selection of toys from the box and when it was playing with the toys, the puppet displayed no emotional affect toward the toys. The experimenter then put away the puppet, had the child help put the five toys the puppet had selected back in the box, and put the box away.

2.1.3. Test phase

The experimenter then brought out two small, clear trays, each containing five of one of the two kinds of toys that had been in the box: the uncommon (17%) toy chosen by the puppet and the more common (83%) toy that the puppet had not chosen. The child then participated in two test procedures, presented in a counterbalanced order. A preference inference task investigated whether children used the statistical non-randomness of the puppet's choices to make inferences about which toy the puppet likes, and a happiness inference task explored whether children can also use this information to make broader inferences about the puppet's emotional reactions to the different toys by asking children to infer which toy had caused the puppet to react happily upon seeing it.

In the preference inference task, the experimenter brought out the gorilla puppet and said to the child, "Look, Gabby's back! And she wants to play with some toys again. Can you give her the one she likes?" If the child hesitated on selecting a toy from one of the trays to give to the puppet, or if the child asked the experimenter which toy to give the puppet, the experimenter simply replied, "Can you give her one?" The puppet then briefly played with the toy the child had given it and put it back in the appropriate tray.

For the happiness inference task, the experimenter brought out a small, round, opaque box filled with tissue paper. The experimenter told the child, "This box has something inside it for Gabby." She then pointed to each of the two trays containing the two kinds of toys from the box in the demonstration phase and said, "It might have a [toy 1] like this or it might have a [toy 2] like this. So look! Here's Gabby! Let's see what she thinks about what's in here." The experimenter always pointed to the toy on her right first, and which kind of toys were on the right was randomized between subjects and between blocks. The puppet looked inside the box and responded happily, clapping its hands and nodding its head. The experimenter then asked the child, "What do you think is in here that makes her happy?" If children did not answer immediately, the experimenter pointed to each tray and asked, "Do you think it's one of these or one of these?"

After the preference and happiness inference tasks, the experimenter put the puppet away. To determine children's own preferences for the toys, the experimenter pointed to the two trays and asked the child, "Can you tell me, which one of these two kinds of toys do *you* like best?" Finally, in order to test whether children accurately remembered the puppet's choices, the experimenter brought out the box from the demonstration phase and asked children to recall which kind of toy the puppet had chosen from that box.

After the participant had answered all the test questions, the experimenter brought out a new box containing the other set of toys that had not been used in the first trial (basketballs and baseballs or ducks and frogs) and repeated the demonstration phase, starting from the introduction of the demonstration box, and the test phase. Individual children received the same framing information (emotion or action) in both trials. In addition, if a child had received the preference inference question first in the first trial, he or she received the happiness inference question first in the second trial, and vice versa. On each of the two trials, children were required to select one of the two kinds of toys as their answer to each of the test questions, and so for each test question children could show three patterns of results: selecting the toy the puppet had chosen from the demonstration box (the target toy) in both trials, selecting the other toy that the puppet had not chosen from the box (the non-target toy) in both trials, or selecting the target toy in one trial and the non-target toy in the other trial; thus, for each test question, each child could select the target toy on zero, one, or two of the two trials. At the end of the second trial, all children received stickers and a certificate of completion.

2.2. Results

In order to determine whether participants used the puppet's choices to make inferences about its mental states, we analyzed the number of trials (out of two) that children in each framing condition selected the toy the puppet had chosen from the demonstration box (the target toy) as their response to the test questions. Preliminary analyses showed no significant effects of participant age, sex, or location tested, so these factors were not included in the following analyses. There were also no effects of counterbalanced factors (toy set used, toy set seen first, question order, or target toy side). Children remembered which toy the puppet had chosen from the box on 86% of trials. All trials are included in the main analyses, and results when excluding children who performed at chance or lower on the memory questions (that is, got the memory question correct on one or zero out of two trials) are also provided. In this and all following studies, all analyses are two-tailed with alpha = .05.

2.2.1. Preference inference

To compare children's performance from chance (selecting the target toy on one out of two trials), we performed a two-tailed binomial sign test comparing the number of children who performed above chance (selected the target toy as the toy the puppet liked on both trials) compared to the number who performed below chance (chose the non-target toy on both trials) – if children were choosing between the two toys at random, we would expect an approximately equal number of children in each group. These analyses revealed that children in the emotion framing condition performed significantly above chance (12/16 children above chance versus 3/16 below chance, $p = .035$), while participants in the action framing condition did not (3/16 children both above and below chance, $p = 1$). To test whether responses differed

depending on framing condition, we performed a Fisher's exact test comparing the number of participants who selected the target toy as the toy the puppet liked in zero, one, or two of the two trials by condition. The results showed that the number of children with each response pattern was significantly dependent on their framing condition ($p < .001$). As shown in Fig. 2, children were more likely to predict that the puppet liked the kind of toy it had chosen from the box when they received emotion framing than when they received action framing. When excluding children who answered the memory question incorrectly on at least one of two trials (seven children total – two in the emotion framing condition and three in the action framing condition who failed on one trial, and one child in each condition who failed on both trials), the results do not change (comparison from chance: emotion framing, $p = .039$, action framing, $p = 1$, sign tests; comparison between conditions: $p = .013$, Fisher's exact test).

2.2.2. Happiness inference

When asked to infer which toy was likely to have caused the puppet's happy reaction upon seeing it in a box, children performed marginally above chance in the emotion framing condition (7/16 children above chance versus 1/16 below chance, $p = .070$, sign test), while children in the action framing condition showed no significant differences from chance in their response patterns (5/16 above chance versus 2/16 below, $p = .45$, sign test.). A Fisher's exact test showed that the difference between the conditions in the number of children who made target selections on zero, one, or two trials was not statistically reliable ($p = .78$). When excluding the seven children who failed at least one memory test, the emotion framing condition becomes significantly different from chance ($p = .031$, sign test), while the action framing condition remains no different from chance ($p = .13$, sign test). The comparison between conditions remains non-significant ($p = .69$, Fisher's exact test).

2.2.3. Child's own preference

When selecting which toy they preferred, children were not influenced by the toy the puppet had chosen during the demonstration phase; that is, children's target choices on the question about their own preferences did not differ from chance ($p = .18$, sign test). In addition, children did not egocentrically assume that the puppet would share their preferences – children did not deviate from chance in selecting the same toy as both the toy they liked and the toy the puppet liked ($p = .14$, sign test), or in selecting the same toy as both the toy they liked and the toy that had made the puppet happy ($p = .33$, sign test). When excluding memory failures, the results do not change (all ps = *n.s.*).

2.3. Discussion

In this study, we found that preschool children only used the statistical non-randomness of an agent's choices to make inferences about the agent's preferences when the verbal framework through which they could interpret the agent's actions was specifically related to mental states. While the emotion framing condition replicates previous work showing that preschoolers can infer preferences from statistically non-random choices when these choices are placed in an explicitly preferential context (Kushnir et al., 2010; Ma & Xu, 2011), the action framing condition shows that without a preference-related framework, children do not make these kind of inferences, despite receiving identical intentional and statistical information as children who did receive a preference framework. Mental state framing information also showed some evidence of supporting children's broader inferences about an agent's future emotional reactions, particularly

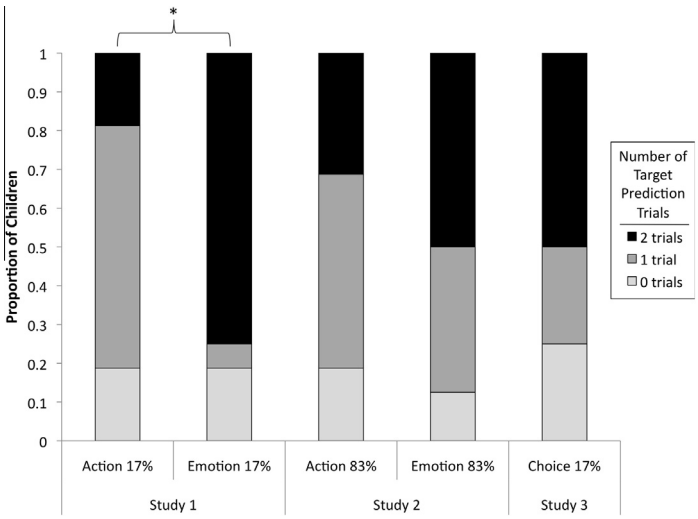

Fig. 2. Which toy does he like? This figure shows the proportion of children who predicted that the puppet liked the toy it had chosen from the box (the target toy) in zero, one, or two out of two trials when the puppet's actions were framed either in terms of emotions, actions, or choices and when the target toy made up either 17% or 83% of toys in the box. Asterisks represent response patterns that are significantly different from each other (Fisher's exact test, $p < .05$).

when considering only children who performed well on the memory questions. However, these results were less strong, suggesting that reasoning about the emotional implications of preferences may be particularly difficult for children (see Section 5 – General discussion for further consideration of this finding). The results of this study suggest that even when provided with systematic behavioral information, children rely on specific contextual information to inform their inferences about which mental states may be driving the behavior.

Study 1 demonstrates that verbal framing information specifically about mental states supports children's mental state inferences when given strong statistical information; however, an open question is whether in some cases, a strong mentalistic verbal framework may also support children's inferences about mental states even in the absence of reliable statistical information. Previous work has shown that 2-year-old children require both compelling statistical evidence and a verbal mental state framework in order to make systematic mental state inferences from non-random choices (Kushnir et al., 2010; Ma & Xu, 2011). For preschool children, however, strong mentalistic verbal framework information may be enough to drive mental state inferences even in ambiguous statistical situations. While previous work has shown that mentalistic verbal framing is not sufficient for preschool children to infer mental states when an agent has not truly chosen between objects (that is, when an agent selects a toy out of a box containing only that kind of toy; Kushnir et al., 2010), the question remains how verbal framing affects children's mentalistic interpretations of choice behavior that is both probabilistic and statistically ambiguous. This question is investigated in Study 2.

3. Study 2

In order to determine the role of verbal framing information about mental states on children's inferences when statistical information is ambiguous, we showed children the same demonstration as in Study 1 of a puppet choosing five identical toys from a box containing 17% of one kind of toy and 83% of another, except in this study the puppet chose the more common (83%) toy from the box. As in Study 1, the puppet's actions were verbally framed either in terms of preference-related emotions or in terms of non-mentalistic actions. In both conditions, as in Study 1, the puppet's actions were intentional and probabilistic, but unlike in Study 1, these actions could not be statistically differentiated from choosing toys randomly, providing insight into the effect verbal framing information may have on children's ability to infer mental states from statistically ambiguous behavior.

3.1. Methods

3.1.1. Participants

Participants were 32 preschool-age children (range = 3 years, 6 months – 4 years, 0 months; mean = 3 years, 9 months; 16 females). Two additional children were tested but excluded because of child interference during the puppet's choice task (1) or failure to complete the study (1). As in Study 1, participants were recruited from a large Midwestern city and tested either in an on-campus laboratory or at a local science museum. The sample of children was 59% Caucasian, 16% multiracial, 9% African-American, 6% Hispanic, 6% Asian, and 3% unreported.

3.1.2. Procedure

The procedure and materials used in Study 2 were identical to those of Study 1, except that the puppet chose the more common (83%) toy from the demonstration box instead of the less common (17%) toy as in Study 1 (see Fig. 1). The emotion framing and action framing provided, as well as all other demonstrations, test questions, counterbalances, and materials were the same as were used in Study 1.

3.2. Results

As in Study 1, preliminary analyses showed no effects of participant age, sex, or location tested, so these factors were not included in the following analyses. There were also no effects of counterbalanced factors, except where noted below. Children remembered which toy the puppet had chosen from the demonstration box on 91% of trials.

3.2.1. Preference inference

In both framing conditions, children's selections of the target toy as the toy the puppet liked did not differ from chance (emotion framing: 8/16 children above chance versus 2/16 below chance, $p = .109$, sign test; action framing: 5/16 above chance versus 3/16 below chance, $p = .73$, sign test). In addition, a Fisher's exact test showed that the framing condition (emotion or action) did not significantly affect the number of participants who selected the target toy as the toy the puppet liked in zero, one, or two trials ($p = .63$). When excluding children who failed the memory test on at least one of two trials (five children total – four who failed on one trial and one who failed on both trials, all in the emotion framing condition), the emotion framing condition becomes significantly different from chance ($p = .008$, sign test; action framing, $p = .73$), and the conditions become marginally different from one another ($p = .09$, Fisher's exact test).

In order to investigate the relative effects of framing condition and the target toy's proportion in the box on children's inferences, the results from Study 2 were combined with the results from Study 1. We performed an analysis on a cumulative logistic regression model with the number of trials children selected the target toy as the toy the puppet liked (zero, one, or two) as an ordinal dependent variable and framing condition (emotion versus action), target toy proportion (17% versus 83%), and the interaction between framing condition and target toy proportion as factors. The results showed a significant main effect of framing condition on children's inferences (Likelihood ratio [LR] $\chi^2(1) = 6.76$, $p = .009$), but no effect of target toy proportion and no interaction effect (proportion effect LR $\chi^2(1) = .36$, $p = .55$; interaction term LR

$\chi^2(1) = 1.50$, $p = .22$).[1] The results of the cumulative logistic regression analysis remain unchanged when excluding the 12 children across both studies who failed the memory test on at least one trial (significant effect of framing, LR $\chi^2(1) = 10.22$, $p = .001$; no effect of proportion or interaction between framing and proportion, $ps > .7$).

3.2.2. Happiness inference

Children in both framing conditions did not significantly differ from chance in their inferences about which toy the puppet had reacted happily to seeing in a box (emotion framing: 4/16 children above chance versus 3/16 below chance, $p = 1$, sign test; action framing: 5/16 above chance versus 2/16 below chance, $p = .45$, sign test). A Fisher's exact test showed that the distribution of participants who selected the target toy as the toy that caused the puppet to react happily upon seeing it in a box in zero, one, or two trials did not differ by framing condition ($p = 1$). When excluding memory failures, the results do not change (all *ps* = *n.s.*).

Although no other counterbalanced factors affected children's inferences, binomial logistic regression analyses showed a significant effect of the target toy on both children's preference inferences and happiness inferences (preference inference: LR $\chi^2(3) = 13.08$, $p = .004$; happiness inference: LR $\chi^2(3) = 8.04$, $p = .045$). This was largely driven by children being more likely to select the target toy as the toy the puppet liked or the toy that would make the puppet happy when the target toy was a basketball than when the target toy was a baseball (for preference inference, 14/16 target selections when target was basketball versus 5/16 when target was baseball; for happiness inference, 13/16 target selections when target was basketball versus 6/16 when target was baseball). In both cases, this effect did not interact with the framing condition (interaction term LR $\chi^2(3) = 2.28$ for preference and LR $\chi^2(3) = 4.94$ for happiness, both *ps* = *n.s.*). The target toy was fully counterbalanced between participants and conditions and no significant target toy effects were found in Study 1 or Study 3, both of which also used these toys, so it is unlikely this result is meaningful; however, this effect may merit further consideration in future work using these toys.

3.2.3. Child's own preference

As in Study 1, children did not systematically prefer the toy the puppet had chosen ($p = 1$, sign test from chance), and they did not systematically predict that the puppet would like or react happily to the child's preferred toy (liking inference: $p = .19$, happiness inference: $p = .36$, sign tests). The results do not change when excluding memory failures (all *ps* = *n.s.*).

3.3. Discussion

Together with the results of Study 1, the results of Study 2 suggest that verbal framing information may provide some support for children's preference inferences even when the behavioral information is statistically ambiguous. When combining the results from Studies 1 and 2, the analysis showed a significant effect of having emotion framing compared to action framing on children's preference inferences, but no significant effect of whether the agent had chosen the more or less common toy from the box and no significant interaction between framing information and more or less common toy selection, indicating that even though the agent's toy selections in Study 2 were not statistically differentiable from random selections, children who received emotion framing still used this behavioral information to make some inferences about the agent's preferences. Although null effects must always be interpreted with caution, and the generally weaker pattern of results in Study 2 compared to Study 1 indicate that the role of strong statistical information cannot be wholly discounted, nevertheless these findings suggest that in some cases, information provided by a strong, relevant framework may be more helpful to children's inferences than ambiguous information provided by an agent's behavior. Interestingly, in this study the preference-relevant verbal framing showed no evidence of supporting children's inferences about which toy had caused the agent to react happily upon seeing it in a box, suggesting that these more difficult emotional inferences about preferences may require additional statistical support (see Section 5 – General discussion). Although strong statistical evidence may be necessary and beneficial for supporting many inferences, these findings highlight the role that explicit verbal information may play in helping children make sense of ambiguous behavioral information.

While the results from Studies 1 and 2 showed that children were better at using an agent's choices to infer her preferences when these choices were placed in an emotion-related framework than when they were described only in terms of actions, the question remains whether children's preference inferences were truly improved by the emotion framework or whether instead, the action framework may have diminished children's performance. Although children received many cues to the agent's intentionality in these studies, it is possible that the phrasing of the action framing conditions, through referring to which toys the puppet "gets," may have biased children in these conditions to view the puppet's actions as less intentional than children in the emotion framing conditions, causing children to be less likely to attribute the puppet's actions to its mental states. Therefore, in Study 3 we investigated the effect of a verbal framework that was intentional but did not contain information about specific mental states.

4. Study 3

In the action framing conditions of Studies 1 and 2, one possible explanation for children's failure to ascribe preferences to the puppet based on its choices is that the wording of this condition may have caused children to perceive the puppet's actions as less intentional. In order to rule out this possibility, we gave children a verbal framework that

[1] The results are similar when results are analyzed by trial with a repeated measures binary logistic regression model: main effect of framing $\chi^2(1) = 5.38$, $p = .020$; no effect of proportion and no proportion * framing interaction (*ps* > .3).

highlighted the intentionality of the puppet's choices but did not speak to the puppet's specific mental states: a *choice* framing. Describing the puppet's actions as choices emphasized that the puppet was intentionally and actively selecting certain toys over others, thus avoiding a possible passive interpretation of the word "gets." With this framing, as with the action framing of Studies 1 and 2, the possible mental state driving the puppet to choose these toys was not explicitly provided. By studying children's inferences about the puppet's statistically non-random behavior when given choice framing, we can determine the role of another kind of framework on children's inferences about mental states.

4.1. Methods

4.1.1. Participants

Participants were 16 preschool-age children (range = 3 years, 6 months – 3 years, 10 months; mean = 3 years, 8 months; 8 females). One additional child was tested but excluded because of child interference during the puppet's choices. Participants were recruited from a large Midwestern city and tested in an on-campus laboratory. Of children sampled, 50% were Caucasian, 19% were African-American, 19% were multiracial, and 13% were Hispanic.

4.1.2. Procedure

The procedure and materials of Study 3 were identical to that of Study 1, except that instead of receiving emotion or action framing, all children received choice framing. When introduced to the puppet, children were told, "This is Gabby and she's going to play with some toys today. Some toys she always *chooses*, and other toys she doesn't *choose*. Do you think you can help me watch and see what she *chooses*?" In each block, before the puppet chose five of the less common (17%) toys from the box, the experimenter said to the puppet, "I have some toys here and I think some you might *choose* and I think some you might not *choose*." After the puppet's toy selection, the experimenter said, "Gabby, is that what you *chose*?" All other demonstrations, test questions, counterbalances, and materials were the same as were used in Study 1.

4.2. Results

As in Studies 1 and 2, preliminary analyses showed no effects of participant age, sex, or location tested, as well as no effects of counterbalanced factors, so these factors were not included in the following analyses. Children correctly remembered which toy the puppet had chosen on 84% of trials.

4.2.1. Preference inference

When asked which toy the puppet likes, children selected the toy the puppet had selected from the box (the target toy) no more often than would be expected by chance (8/16 children above chance versus 4/16 below chance, $p = .39$, sign test). Furthermore, the number of children who selected the target toy as the toy the puppet liked in zero, one, or two trials in this study was not significantly different from either condition of Study 1, which provided children with identical statistical information as in this study (versus Study 1 emotion: $p = .57$, versus Study 1 action: $p = .16$, Fisher's exact tests, adjusted for two comparisons using Bonferroni method). The results do not change when excluding children who failed at least one memory question (four children total – three who failed on one trial and one who failed on both trials; comparison from chance: $p = .11$, sign test; comparison to Study 1: versus emotion framing $p = 1$, versus action framing $p = .15$, Bonferroni-adjusted Fisher's exact tests).

4.2.2. Happiness inference

Children's inferences about which toy caused the puppet to react happily upon seeing it in a box did not differ from chance (6/16 children above chance versus 2/16 below chance, $p = .29$, sign test). In addition, the number of children who selected the target toy as the toy that caused the puppet to be happy after seeing it in a box on zero, one, or two trials was not significantly different from either framing condition in Study 1 (Bonferroni-adjusted Fisher's exact tests, both $ps = 1$). Again, the results do not change when excluding the four children who performed below chance on the memory questions (comparison from chance: $p = .45$, sign test; comparison to Study 1: versus emotion framing $p = 1$, versus action framing $p = .72$, Bonferroni-adjusted Fisher's exact tests).

4.2.3. Child's own preference

The number of children who preferred the target toy was marginally different from chance; however, children were more likely to perform *below* chance than above chance – that is, children were marginally more likely to prefer the toy the puppet had *not* chosen from the box than the toy the puppet had chosen (1/16 children with target selections above chance versus 7/16 below chance, $p = .070$, sign test). As in Studies 1 and 2, children did not systematically predict that the puppet liked or would be happy about the child's preferred toy (preference inference $p = 1$, happiness inference $p = .75$, sign tests). When excluding memory failures, all results are nonsignificant (all $ps = n.s.$).

4.3. Discussion

When preschool children were given verbal framework information that drew attention to the intentionality of the agent's choices but was not informative about any specific mental state, they did not use the agent's statistically non-random choices to make systematic inferences about the agent's preferences. These results show that the findings from Studies 1 and 2 were likely not due to the action framing biasing children to view the agent's actions as non-intentional; rather, these results underscore the importance of having a framework that is specifically relevant to the task at hand. Although referring to an agent's choices ostensibly could be related to preferences, this study suggests that preschool children's understanding of the relationship between choices and preferences is not strong enough to significantly influence their inferences in this case. Because the results do not significantly differ from either framing condition in Study 1, future research

is needed to determine whether children's inferences after receiving choice framing may pattern more closely with emotion framing as compared to action framing. Nevertheless, these findings demonstrate that simply highlighting that the agent's choices were intentional – a task also accomplished by the agent selecting items in a statistically non-random way – was not sufficient for children to make systematic mental state inferences based on this behavior.

5. General discussion

The current findings expose the vital role that explicit, relevant verbal framework information can play in driving children's inferences from statistical information, particularly when using statistically non-random behavior to infer others' mental states. When children observed an agent choose certain toys statistically more often than chance, children only used this information to predict the agent's preferences when the agent's choices were framed with specifically preference-related verbal information. Children did not systematically infer the agent's preferences when the agent's behavior was framed as actions ("Gabby gets some toys") or as choices ("Gabby chooses some toys"), rather than mental states. Furthermore, mental state framing showed some influence on children's mental state inferences even when the statistical regularity of the agent's choices was uninformative about her preferences.

These results suggest that an important component of successful Bayesian reasoning – selecting appropriate hypotheses to consider before observing statistical information and evaluating these hypotheses after observing the data – may be difficult for young children, especially in the context of mental state reasoning. In the current studies, the puppet could plausibly have had any number of reasons for selecting five certain toys from the box, one of which being because it likes them. However, in the absence of a specific prior theory about the puppet's preferences, it is equally plausible that the puppet may have been removing unwanted toys from the box, organizing the toys by color, or even choosing toys for someone else to play with. The current findings show that preschool children were unable to spontaneously posit preferences as the most likely explanation for the puppet's choices without being provided with verbal scaffolding information that was specifically relevant to preferences. Thus, having a specific framework theory may be vital for restricting the number of possible explanations for a statistical pattern, in this case enabling children to use an agent's behavior to make appropriate inferences about the agent's mental states. With a more constrained prior hypothesis about which mental states were most likely to have given rise to the agent's actions, children can focus on whether the provided statistical information supports or refutes that possible hypothesis, an analytical ability young children have demonstrated in previous research (Kushnir et al., 2010; Lucas et al., 2014; Ma & Xu, 2011), and then use this information to formulate new hypotheses. The current findings demonstrate not only that changing children's prior theories can influence how statistical data are interpreted, but also that without a prior theory that is specifically relevant to the task at hand, preschool children may have difficulty drawing any substantial conclusions about data.

The combined results of Study 1 and Study 2 show that specific framework theories can also help support some inferences even from statistically ambiguous behavior, raising the possibility that framing information may in some situations have stronger effects on children's inferences than ambiguous statistical information. It has been well documented that an especially strong framework or prior belief can bias adults not only to specifically search out data that confirm this belief, but also to interpret ambiguous data as providing support for their beliefs (e.g., Lord, Ross, & Lepper, 1979; see Rabin & Schrag, 1999 for a review). The results of Study 2 suggest that in some cases, children may be subject to this same confirmation bias. When preschool children were given strong, explicit verbal cues that the puppet's choices related to its preferences, they may have been biased to seek out evidence that helps confirm their belief that the puppet prefers certain toys, which may have caused children to view even the weak data presented in Study 2 as somewhat informative about the puppet's preferences. Thus, specific framework theories may be important not only for changing how statistical data are interpreted, but also for influencing which information is regarded as useful data in the first place. Future research can help determine the extent to which children's inferences about ambiguous statistical information may be affected by specific framing information.

One potentially surprising finding from the current studies is that while children were able to use the emotion framework to help them determine which toy the puppet liked, this information was less helpful in helping children make broader inferences about the puppet's future emotional reactions (that is, inferring which toy had caused the puppet to react happily upon seeing that toy in a box), despite the emotion framing consisting of the experimenter specifically telling children that some toys made the puppet happy. One possible explanation for this result is that the happiness question required a more difficult inference than the preference question. To answer the preference inference question, children simply had to give the puppet the toy it liked, but to correctly answer the happiness inference question children had to also understand that a certain toy had caused the puppet's happy reaction and mentally represent which toy might have been in the box to draw this reaction. Despite the difficulty of this task, children in Study 1 who both received emotion framing and observed the puppet choosing toys statistically non-randomly performed marginally above chance, and significantly above chance when considering only children who performed well on the memory task, suggesting that for more inferentially opaque tasks, relevant verbal framing information and strong statistical information may work together to support children's reasoning. For instance, children in Study 1's emotion framing condition may have been driven by the statistical non-randomness of the puppet's choices and the emotional content in the verbal framing information to search for a deeper

relationship between the puppet's preferences and its emotional reactions (see Griffiths & Tenenbaum, 2007). Future research could further investigate the factors that support children in making these more complex inferences.

5.1. Novel inferences from statistical information

The current findings suggest that in the absence of a strong, relevant mental state framework, preschool children may have difficulty formulating novel psychological explanations for statistical patterns in behavior. These results may be somewhat surprising because previous research indicates that children can make some novel inferences from statistical information when they are reasoning about the physical and biological causes of a situation as opposed to psychological causes of behavior. Work on children's physical causal reasoning has shown that preschool children notice evidence that is inconsistent with their prior causal beliefs and are driven to formulate explanations for it (Legare, 2011; Legare, Gelman, & Wellman, 2010). With enough surprising or inconsistent evidence, children can overcome strongly held physical or biological beliefs, such as that items must touch in order to cause physical outcomes, and formulate alternate explanations that fit with the observed patterns of evidence, such as that causality may sometimes work at a distance (Kushnir & Gopnik, 2007; see also Schulz, Bonawitz, & Griffiths, 2007; Schulz & Gopnik, 2004). Preschool children can also infer the presence of "unobserved causes" – potential causes that children are aware of but have not seen directly operate on a system (Gopnik et al., 2004; Schulz, Goodman, Tenenbaum, & Jenkins, 2008; Schulz & Sommerville, 2006). In the current studies, however, preschool children were unable to use anomalous or surprising behavioral information to formulate an unobserved psychological mechanism that may differ from their previous understanding of mental states without support from specific, mentalistic verbal framework information.

One likely explanation for the results of the current studies is that novel, abstract, conceptual inferences about how mental states drive behavior may be more difficult than the inferences children were required to make in previous work on physical and biological causes. Preschool children often have difficulty reasoning about others' mental states and the behavior these mental states engender (see Wellman & Liu, 2004), and so a novel inference about a specific mental state causing a specific pattern of behavior may be especially challenging for young children to comprehend. Furthermore, although the inferences children made in previous studies were novel in that children had to infer causal mechanisms that they had not considered before viewing the evidence, children in these studies also generally had the pieces of the causal mechanism in place and only had to infer the novel causal links between them. For example, while children were able to infer that a flashlight could cause a box to light up in a certain way or that a certain psychological state could have biological effects, children had been provided with information about both the flashlight and the psychological states in the course of the experiment (Schulz & Sommerville, 2006; Schulz et al., 2007). When inferring the novel psychological causes of behavior, the psychological cause is often not directly observable, and thus children may have difficulty reasoning about a kind of mental state they know little about.

Although making inferences about mental states may be difficult for young children, the current studies demonstrate that verbal information can play a vital role in helping shape the set of potential psychological explanations children posit for observed behavior. Many researchers have argued for the importance of language in children's conceptual development (e.g., Astington & Baird, 2005; Gelman, 2003; Gopnik & Meltzoff, 1997; among many others). In this case, through attending to verbal information and other contextual factors, children can become more familiar with the many different mental states that drive behavior. Contextual information such as verbal framing may be essential to helping children appropriately analyze statistical information and may even be the driving force behind children's ability to derive novel conceptual insights from patterns in the world.

5.2. What kind of framing matters?

Although this study focused on the effects of verbal framing information, an important open question is whether other kinds of framing may be equally important for driving children's inferences from behavioral data. One such framework that may affect children's mental state inferences is nonverbal affective information. Evidence suggests that when making physical causal inferences, young children are sensitive to affective information, making more accurate causal inferences when an experimenter demonstrates surprise rather than neutral affect at causal failures (Sobel, Sommerville, Travers, Blumenthal, & Stoddard, 2009). In a similar way, an agent's behavior being accompanied by relevant affective information may help guide children to more accurately assess the likely mental state driving this behavior. For instance, if an agent reacts with positive affect to each toy she pulls out of a box, children may infer that she likes that kind of toy even without any verbal framing information. In contrast, if an agent has a strong negative reaction to each toy chosen from a box, children may infer that the agent has a dispreference for that item, even if she had chosen those items non-randomly. As with verbal framing information, affective information may give children a relevant framework through which to interpret an agent's behavior. Further research is needed to evaluate this possibility.

Other kinds of contextual information present in a situation may affect children's inferences about statistical information as well. For instance, a great deal of research has shown that evidence provided through pedagogical instruction leads to different inferences about a variety of systems than evidence provided non-pedagogically (e.g., Bonawitz et al., 2011; Buchsbaum, Gopnik, Griffiths, & Shafto, 2011; Butler & Markman, 2012b, 2014; Shafto, Goodman, & Griffiths, 2014). Similarly, young children make different inferences about statistical evidence when it is generated by a knowledgeable teacher rather than by themselves (Xu & Tenenbaum, 2007) and when the evidence is generated intentionally rather than accidentally

(Gweon et al., 2010). Given these findings, it seems possible that pedagogical demonstrations may have similar effects on children's inferences about others' mental states when interpreting behavioral data. However, in the current study the behavioral information was presented pedagogically and intentionally in all cases, yet children only inferred the agent's preferences when the behavior was verbally framed in terms of the agent's mental states. In this situation, pedagogical demonstrations were not sufficient to guide children's inferences about mental states; instead, the specific mental state content of the verbal framing appears to have been the driving force for supporting children's inferences. Although in some cases, attending to broader contextual factors such as the pedagogical intent of an agent may provide an additional kind of support for children's inferences about statistical data, in other situations children may need a more explicit and specific framework in order to accurately interpret the data.

5.3. Framing and folk theories

In addition to external factors such as verbal, affective, and contextual information, an observer's own internal theories and expectations about how the world works, often called folk theories, may also shape his or her interpretations of statistical information. When interpreting behavior, these folk theories may function by providing observers with a default explanation of the mental state causing the behavior. For example, a strong belief that people choose to eat their favorite foods at restaurants may lead you to infer that your dinner partner ordered a salad instead of a burger because she really likes salads, not because she wants to eat healthier food, is a vegetarian, or has heard that the restaurant serves especially delicious salads. Interestingly, evidence suggests that young children may not employ a strong default explanation or folk theory when reasoning about other people's preferences. Previous research has shown that young children do not have a strong folk understanding of how preferences work and often fail to view preferences as stable personality traits that can predictably drive behavior (Kalish, 2002; Kalish & Shiverick, 2004). Furthermore, in the current study children failed to make systematic preference inferences in the absence of specific preference-related verbal framing information, suggesting that preschool children may not have a sufficiently developed folk understanding of preferences to spontaneously assume that preferences may be the best explanation for a character's choices.

Although adults' understanding of preferences is more sophisticated than those of children (Kalish, 2002; Kalish & Shiverick, 2004), these folk theories are culturally variable. While North American adults hold strong beliefs that people are motivated to express their preferences through their choices and that choices reflect a person's preferences, Indian adults are less likely to see preferences as a primary motivator of a people's choices (Savani, Markus, & Conner, 2008). This suggests that the development of folk theories of preferences may be highly affected by experience, raising the possibility that an important function of external verbal framework information may be to help shape people's internal folk theories about the world. When people experience many instances of a certain behavior being explicitly framed to allow for specific mental state explanations, over time they may develop default expectations about which mental states drive certain behaviors. In this way, explicit framework theories can allow people to generalize their experiences in order to make accurate inferences even in novel situations.

An interesting implication of folk theories of preferences developing throughout childhood is that as default mental state explanations become increasingly established and salient, the importance of explicit, highly relevant verbal framing information may decrease. Because young children have a weaker folk understanding of preferences than adults (Kalish, 2002; Kalish & Shiverick, 2004), mentalistic verbal framing may play an especially important role in guiding children's mental state inferences by explicitly drawing the connection between an agent's choices and his or her preferences, making clear, unambiguous mental state framing information necessary for children to be able to infer preferences from an agent's choices. In contrast, adults and even older children may not require such specific verbal framework information to make mental state inferences from behavior because their inferences may be driven primarily by strong default expectations about the mental states that often drive certain behavior. Furthermore, certain verbal framing information may be interpreted as more clearly related to preferences and other mental states with age, such as the choice framing used in Study 3. Future research can help determine how and whether the importance of specifically relevant verbal framework information for supporting mental state inferences changes throughout development.

5.4. Conclusions

In this work, we found that preschool children only used statistically non-random choices to infer preferences in the presence of a specific mentalistic verbal framework, suggesting that having a specifically relevant framework theory may be vital for allowing young children to accurately interpret statistical information. The current findings demonstrate that while preschool children are able to reason intelligently about how behavioral information informs a specific hypothesis about the psychological cause of the behavior, they require additional support to determine which psychological causes may be possible and plausible. Our findings also highlight that when developing an understanding of others' minds, young children integrate information from many different sources, including both statistical patterns in behavior and verbal information provided by others.

Acknowledgements

A special thank you to the children and families who participated in this research. We would also like to thank Rebecca Schmidt, Anna Pfautz, and the other members of the Infant Learning and Development Laboratory at the University of Chicago for their assistance with this project, as well as the Museum of Science and Industry, Chicago for testing space. We are grateful to Tamar Kushnir and two

anonymous reviewers for their helpful comments and suggestions on an earlier version of this manuscript. This work was supported by an NICHD Grant (R01 HD 035707) to the second author.

References

Astington, J. W., & Baird, J. A. (2005). *Why language matters for theory of mind*. New York, NY: Oxford University Press.

Bonawitz, E. B., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E. S., & Schulz, L. E. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition, 120*(3), 322–330. http://dx.doi.org/10.1016/j.cognition.2010.10.001.

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition, 120*(3), 331–340. http://dx.doi.org/10.1016/j.cognition.2010.12.001.

Butler, L. P., & Markman, E. M. (2012a). Finding the cause: Verbal framing helps children extract causal evidence embedded in a complex scene. *Journal of Cognition and Development, 13*(1), 38–66. http://dx.doi.org/10.1080/15248372.2011.567201.

Butler, L. P., & Markman, E. M. (2012b). Preschoolers use intentional and pedagogical cues to guide inductive inferences and exploration. *Child Development, 83*(4), 1416–1428. http://dx.doi.org/10.1111/j.1467-8624.2012.01775.x.

Butler, L. P., & Markman, E. M. (2014). Preschoolers use pedagogical cues to guide radical reorganization of category knowledge. *Cognition, 130*(1), 116–127. http://dx.doi.org/10.1016/j.cognition.2013.10.002.

Denison, S., Konopczynski, K., Garcia, V., & Xu, F. (2006). Probabilistic reasoning in preschoolers: Random sampling and base rate. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 1216–1221). Florence, KY: Psychology Press.

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. *Developmental Psychology, 49*(2), 243–249. http://dx.doi.org/10.1037/a0028278.

Denison, S., & Xu, F. (2009). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13*(5), 798–803. http://dx.doi.org/10.1111/j.1467-7687.2009.00943.x.

Diesendruck, G., Salzer, S., Kushnir, T., & Xu, F. (2013). When choices aren't personal: The effect of statistical and social cues on children's inferences about the scope of preferences. *Journal of Cognition and Development*. http://dx.doi.org/10.1080/15248372.2013.848870.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press.

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science, 337*(6102), 1623–1627. http://dx.doi.org/10.1126/science.1223416.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111*(1), 3–32. http://dx.doi.org/10.1037/0033-295X.111.1.3.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: The MIT Press.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin, 138*(6), 1085–1108. http://dx.doi.org/10.1037/a0028044.

Griffiths, T. L., Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2011). Bayes and blickets: Effects of knowledge on causal induction in children and adults. *Cognitive Science, 35*(8), 1407–1455. http://dx.doi.org/10.1111/j.1551-6709.2011.01203.x.

Griffiths, T. L., & Tenenbaum, J. B. (2007). From mere coincidences to meaningful discoveries. *Cognition, 103*(2), 180–226. http://dx.doi.org/10.1016/j.cognition.2006.03.004.

Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, 107*(20), 9066–9071. http://dx.doi.org/10.1073/pnas.1003095107.

Kalish, C. W. (2002). Children's predictions of consistency in people's actions. *Cognition, 84*(2002), 237–265. http://dx.doi.org/10.1016/S0010-0277(02)00052-5.

Kalish, C. W., & Shiverick, S. M. (2004). Children's reasoning about norms and traits as motives for behavior. *Cognitive Development, 19*(3), 401–416. http://dx.doi.org/10.1016/j.cogdev.2004.05.004.

Kushnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology, 43*(1), 186–196. http://dx.doi.org/10.1037/0012-1649.43.1.186.

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science, 21*(8), 1134–1140. http://dx.doi.org/10.1177/0956797610376652.

Legare, C. H. (2011). Exploring explanation: Explaining inconsistent evidence informs exploratory, hypothesis-testing behavior in young children. *Child Development, 83*(1), 173–185. http://dx.doi.org/10.1111/j.1467-8624.2011.01691.x.

Legare, C. H., Gelman, S. A., & Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development, 81*(3), 929–944. http://dx.doi.org/10.1111/j.1467-8624.2010.01443.x.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*(11), 2098–2109. http://dx.doi.org/10.1037/0022-3514.37.11.2098.

Lucas, C. G., Griffiths, T. L., Xu, F., Fawcett, C., Gopnik, A., Kushnir, T., et al. (2014). The child as econometrician: A rational model of preference understanding in children. *PLoS ONE, 9*(3), e92160. http://dx.doi.org/10.1371/journal.pone.0092160.s001.

Ma, L., & Xu, F. (2011). Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition, 120*(3), 403–411. http://dx.doi.org/10.1016/j.cognition.2011.02.003.

Ma, L., & Xu, F. (2013). Preverbal infants infer intentional agents from the perception of regularity. *Developmental Psychology, 49*(7), 1330–1337. http://dx.doi.org/10.1037/a0029620.

Muentener, P., & Schulz, L. E. (2012). What doesn't go without saying: Communication, induction, and exploration. *Language Learning and Development, 8*(1), 61–85. http://dx.doi.org/10.1080/15475441.2011.616455.

Newman, G. E., Keil, F. C., Kuhlmeier, V. A., & Wynn, K. (2010). Early understandings of the link between agents and order. *Proceedings of the National Academy of Sciences of the United States of America, 107*(40), 17140–17145. http://dx.doi.org/10.1073/pnas.0914056107.

Perfors, A., Tenenbaum, J. B., Griffiths, T. L., & Xu, F. (2011). A tutorial introduction to Bayesian models of cognitive development. *Cognition, 120*(3), 302–321. http://dx.doi.org/10.1016/j.cognition.2010.11.015.

Rabin, M., & Schrag, J. L. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics, 114*(1), 37–82. http://dx.doi.org/10.1162/003355399555945.

Savani, K., Markus, H. R., & Conner, A. L. (2008). Let your preference be your guide? Preferences and choices are more tightly linked for North Americans than for Indians. *Journal of Personality and Social Psychology, 95*(4), 861–876. http://dx.doi.org/10.1037/a0011618.

Schulz, L. E., Bonawitz, E. B., & Griffiths, T. L. (2007). Can being scared cause tummy aches? Naive theories, ambiguous evidence, and preschoolers' causal inferences. *Developmental Psychology, 43*(5), 1124–1139. http://dx.doi.org/10.1037/0012-1649.43.5.1124.

Schulz, L. E., Goodman, N. D., Tenenbaum, J. B., & Jenkins, A. C. (2008). Going beyond the evidence: Abstract laws and preschoolers' responses to anomalous data. *Cognition, 109*(2), 211–223. http://dx.doi.org/10.1016/j.cognition.2008.07.017.

Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology, 40*(2), 162–176. http://dx.doi.org/10.1037/0012-1649.40.2.162.

Schulz, L. E., & Sommerville, J. A. (2006). God does not play dice: Causal determinism and preschoolers' causal inferences. *Child Development, 77*(2), 427–442. http://dx.doi.org/10.1111/j.1467-8624.2006.00880.x.

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology, 71*(C), 55–89. http://dx.doi.org/10.1016/j.cogpsych.2013.12.004.

Sobel, D. M., Sommerville, J. A., Travers, L. V., Blumenthal, E. J., & Stoddard, E. (2009). The role of probability and intentionality in preschoolers' causal generalizations. *Journal of Cognition and Development, 10*(4), 262–284. http://dx.doi.org/10.1080/15248370903389416.

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science, 28*(3), 303–333. http://dx.doi.org/10.1016/j.cogsci.2003.11.001.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*(2), 523–541. http://dx.doi.org/10.1111/j.1467-8624.2004.00691.x.

Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition, 112*(1), 97–104. http://dx.doi.org/10.1016/j.cognition.2009.04.006.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences, 105*(13), 5012–5015. http://dx.doi.org/10.1073/pnas.0704450105.

Xu, F., & Kushnir, T. (2013). Infants are rational constructivist learners. *Current Directions in Psychological Science, 22*(1), 28–32. http://dx.doi.org/10.1177/0963721412469396.

Xu, F., & Tenenbaum, J. B. (2007). Sensitivity to sampling in Bayesian word learning. *Developmental Science, 10*(3), 288–297. http://dx.doi.org/10.1111/j.1467-7687.2007.00590.x.